Matteo Paggiaro - 10521399 . Applied Statistics 15/6/20

Problema 1

a) Testiamo la Gaussianità di $\underline{X}$:

<u>Ipotesi:</u>

$$\underline{X}_1, \dots, \underline{X}_n \text{ iid } \sim N_p(\underline{\mu}, \Sigma)$$

con $p = 2$

Il <u>Shapiro Test MC</u> restituisce un p-value elevato :

$\boxed{p_{\text{shapiro MC}} \cong 22\%}$ $\Rightarrow$ $\underline{X}$ gaussiano!

Procediamo quindi con il test attraverso la statistica di Hotelling :

$$T_0^2 = n(\bar{\underline{X}} - \underline{\mu}_0)' S^{-1} (\bar{\underline{X}} - \underline{\mu}_0) \overset{H_0}{\sim} \frac{(n-1)p}{n-p} F(p, n-p)$$

con H_0 : $\underline{\mu}_0 = \begin{pmatrix} 50 \\ 50 \end{pmatrix}$

Rifiuteremo H_0 se $T_0^2 > \frac{(n-1)p}{n-p} F_\alpha(p, n-p)$.

In particolare, $(\alpha = 5\%)$ $\begin{cases} T_0^2 \cong 59.6 \\ F_\alpha(\dots) \simeq 6.74 \end{cases}$

$\Rightarrow$ <u>Rifiutiamo H_0!</u> p-value del test $\simeq 9.13 \times 10^{-11}$
↳ basso!

b) Costruiamo la Confidence Region per $\underline{\mu}$:

$$CR_{1-\alpha}(\underline{\mu}) = \left\{ \underline{\eta} \in \mathbb{R}^p : d^2_{S^{-1}}(\bar{\underline{X}}, \underline{\mu}) \le \underbrace{\frac{(n-1)p}{n-p} F_\alpha(p, n-p)}_{r^2_{\text{fisher}}} \right\}$$

<u>Centro:</u> $\bar{\underline{X}} = \begin{pmatrix} 92.07 \\ 111.45 \end{pmatrix}$

<u>Assi:</u> $\underline{e}_1 = \text{eig}(S)_1 = \begin{pmatrix} 0.525 \\ 0.8507 \end{pmatrix}$

$\underline{e}_2 = \text{eig}(S)_2 = \begin{pmatrix} -0.85 \\ 0.525 \end{pmatrix}$ $\perp$

<u>Lunghezza degli assi:</u> $\begin{cases} \sqrt{\lambda_1}/\sqrt{n} \cdot \sqrt{r^2_{\text{fisher}}} \\ \sqrt{\lambda_2}/\sqrt{n} \cdot \sqrt{r^2_{\text{fisher}}} \end{cases}$

$$\begin{cases} \lambda_1, \lambda_2 \rightarrow \text{autovalori di } S \\ e_1, e_2 \rightarrow \text{autovettori di } S \text{ associati a } \lambda_1, \lambda_2. \end{cases}$$

Guardare figura 1 : $\begin{cases} CR(\mu) \text{ ellisse blu} \\ \mu_0 : \text{punto grigio (fuori dall'ellisse)} \end{cases}$

c) Il test sottolinea come $\mu_0 \notin CR_{1-\alpha}(\mu)$.

Se fosse stato contenuto nella confidence region, il test avrebbe avuto un risultato positivo per H_0: avremmo avuto infatti

$$T_0^2 \leq \frac{(n-1)p}{n-p} F_\alpha(p, n-p) \qquad (*)$$

↳ il p-value è l'$\bar{\alpha}$ più piccolo per cui la relazione (*) è valida (in particolare vale l'uguaglianza), ossia il valore minimo in cui $\mu_0 \in CR_{1-\bar{\alpha}}(\mu)$.

d) $$\text{Sim}\,CI_\alpha(\mu_1) = \left[\bar{X}_1 \pm \sqrt{\frac{(n-1)p}{n-p} F_\alpha(p, n-p)} \sqrt{\frac{S_{11}}{n}} \right]$$

$$\text{Sim}\,CI_\alpha(\mu_2) = \left[\bar{X}_2 \pm \sqrt{\quad \dots \quad} \sqrt{\frac{S_{22}}{n}} \right]$$

I risultati sono:

PM2.5 : $\text{Sim}CI(\mu_1) = [77.38, \; 106.78]$

PM10 : $\text{Sim}CI(\mu_2) = [90.32, \; 132.58]$